**设计说明文件**

姓名：叶倩琳

班级：软工1706班

学号：201706061330

**一、实验目的**

编程实现正规式转换到非确定有穷自动机转换的一般算法

**二、实验要求**

1. 用C++ 实现；

2. input.txt描述正规式，需要编程加以解析，该文件规定的格式为：第一行写处正规式的符号集，规定符号集为0-9共10个数字符，第二行写正规式，如：

0123456789

(0|1)\*0.10\* （注：\*表示闭包运算，.表示连接该符号可省略）

3. output.txt为描述非确定有穷自动机的文件，有程序生成，生成时需按如下格式：

第一行写出状态集，用大写字母A-Z表示，如：ABCD

第二行描述自动机的符号集为0-9共10个数字符，如：0123456789

第三行写初态集，如：A

第四行写终态集，如：CD

第五行写出转移函数,如(A,0,B)表示A状态的出边0到达B状态。上例正规式可表为(A,0,A)(A,1,A)(A,0,B)(B,1,C)(C,0,D)(D,0,D)。

4.提供名为retonfa.exe的执行文件实现上述功能，还需提供源文件以及名为设计说明.doc的说明文件。

**三、实验过程**

**（1）流程图**

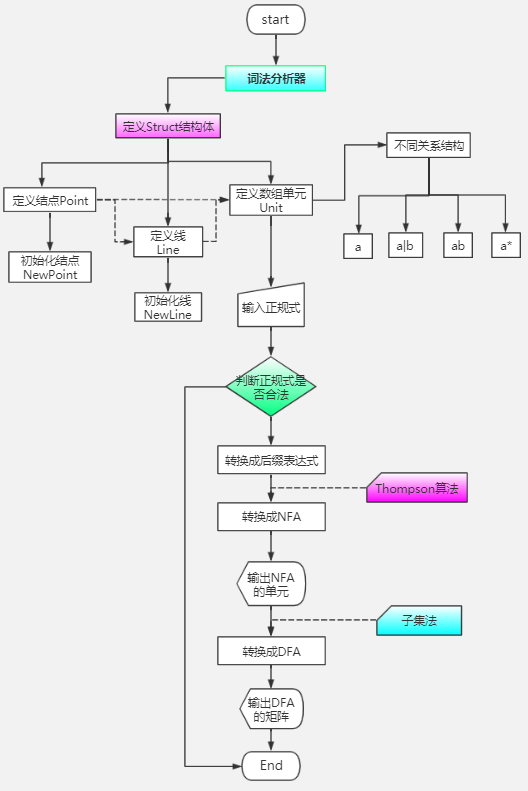


图3-1-1 词法分析器流程图

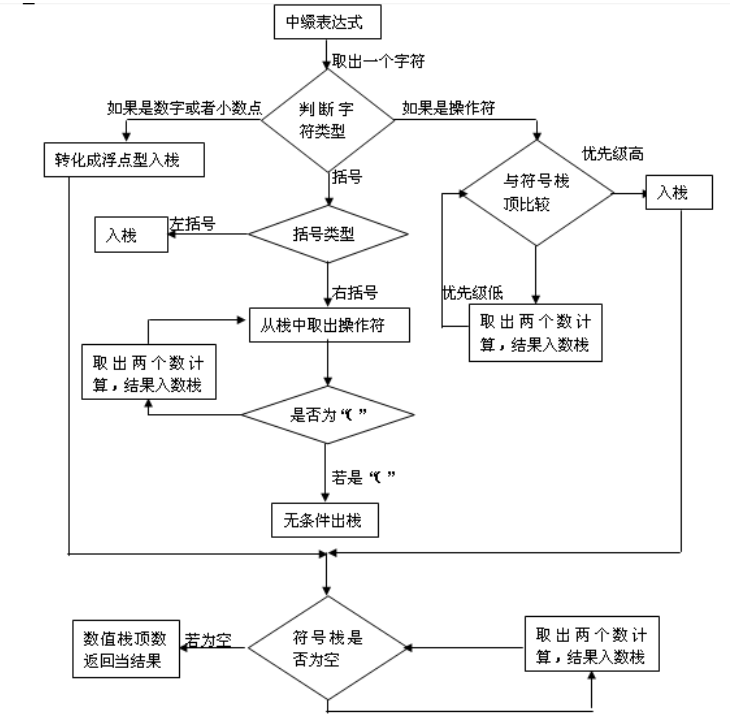


图3-1-2 中缀转后缀流程图

1. **概念介绍**

描述正规语言的形式工具有：3 型（正规）文法、有限自动机和正规表达式，并且上述三者等价。

其中， 一个非确定有限状态自动机 NFA nondeterministic finite automata) 是一个五元组 A = (K,∑,δ,S,Z)。

K有限状态集

∑有限输入符号集

δ转移函数

S非空初态集

Z一个终态集合

1. **程序设计的主要思路**
2. 创建一个结构体Struct edge包含NFA中顶点引出的边。
3. 设置算符优先级函数setprecece，定义 \* > . > |
4. 从文件中读取正规式存入字符串expression中。
5. 加入连接符（.）
6. 中缀转后缀，需要借助栈来实现。中缀表达式转化为后缀表达式，这样就可以去掉'('和')'了，方便构造NFA。

从中缀表达式的左边开始扫描，若遇到运算数时，则直接将其输出（不压入堆栈）。若遇到左括号，则将其压栈。若遇到右括号，表达括号内的中缀表达式已经扫描完毕，这时需将栈顶的运算符依次弹出并输出，直至遇到左括号[左括号弹出但不输出]。若遇到的是运算符，如果该运算符的优先级大于栈顶运算符的优先级时，将其压栈；否则将栈顶运算符弹出并输出，接着和新的栈顶运算符比较，直至运算符大于栈顶云算符为止。最后，若扫描到中缀表达式的末尾，堆栈中还有存留的运算符依次弹出并输出即可。

1. 将后缀表达式转化为NFA。每当遇到操作数时就构造单个符号的NFA对象，压入栈中；每当遇到操作符(\*, ., |)时,就从栈中取出操作数(NFA对象)，进行运算后(\*, |, .)后，将得到的结果NFA压入栈中。最终栈中剩下的那个NFA对象就是与正则表达式等价的NFA对象。其核心算法如下。

首先初始化一个栈h，遍历前面得到的后缀表达式。若遇到非运算符及表达式里面的转移符号时，需要构造一个基本的NFA，由初始状态至终止状态有一条为该转移符号的边，此时仍然需要检查正则表达式的下一个符号，如果不是运算符或者为左括号，此时应该运算栈中添加一个连接运算符，然后将构造的基本NFA添加入NFA栈中，方便以后将基本的NFA进行其他选择，重复，连接运算。

若遇到非运算符时，需要分四种运算符的情况：

1. 若运算符为“）”，要在符号栈中弹出所有符号运算，直到遇到“）”匹配。
2. 若运算符为“（”，直接入栈。
3. 若运算符为“\*”，由于重复符号运算级最高，可直接进行计算，从NFA栈中弹出一张图，然后得到两个未分配的新节点，如图3-4-1-6所示添加4条边，更新起始节点和终止节点后将新的NFA压入NFA栈，运算后检查其后跟随的元素，如果是转移符号或者左括号，则必须要向符号栈中添加连接符号。
4. 若运算符为“+”，应该弹出符号栈中优先级高于它的符号，但“（”不参与弹出，只弹出连接符号和自身“+”符号运算，然后将该符号压入符号栈等候计算。

最后，正则表达式遍历完毕之后，需要弹出所有的符号栈进行计算，最后NFA栈中的唯一NFA就是所求的NFA。

7）获得并打印转移函数。

**（4）绘制ε-NFA**

1）从正规表达式构造等价的ε- NFA（归纳构造过程: Thompson 构造法）

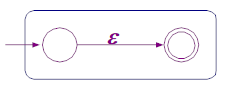


图3-4-1-1 ε的等价NFA



图3-4-1-2 ∅的等价NFA

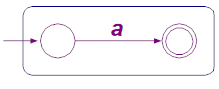


图3-4-1-3 a的等价NFA

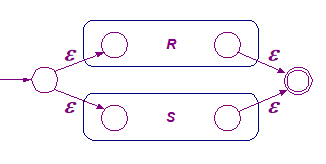


图3-4-1-4 R|S的等价NFA

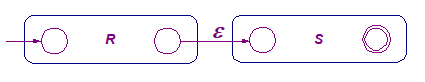


图3-4-1-5 RS的等价NFA

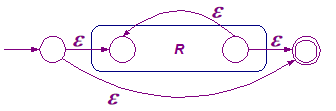


图3-4-1-6 R\*的等价NFA

2）本题中，正规式(0|1)\*0.10\*，首先考虑(0|1)\*的等价NFA。

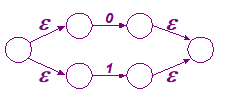


图3-4-2-1 (0|1)的等价NFA

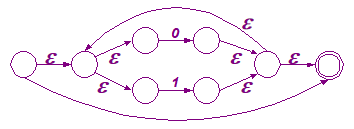


图3-4-2-2 (0|1)\*的等价NFA

其次，考虑0.1.的等价NFA。



图3-4-2-3 0.1.的等价NFA

再考虑0\*的等价NFA。

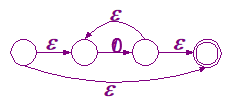


图3-4-2-4 0\*的等价NFA

最后，把上述部分NFA连接起来整合得到等价的ε-NFA。

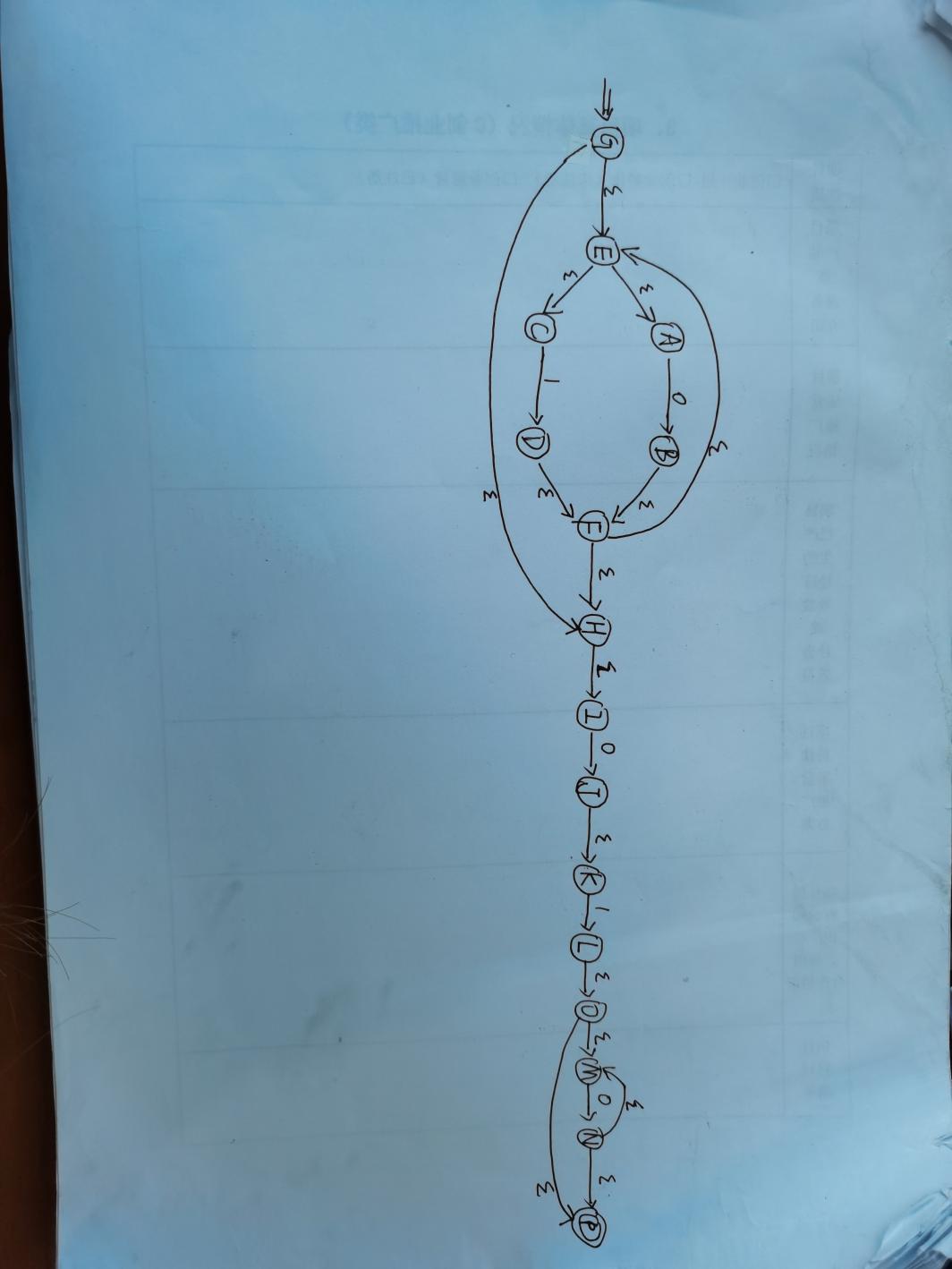


图3-4-2-5 正规式(0|1)\*0.10\*整合的等价NFA

**四、实验结果**

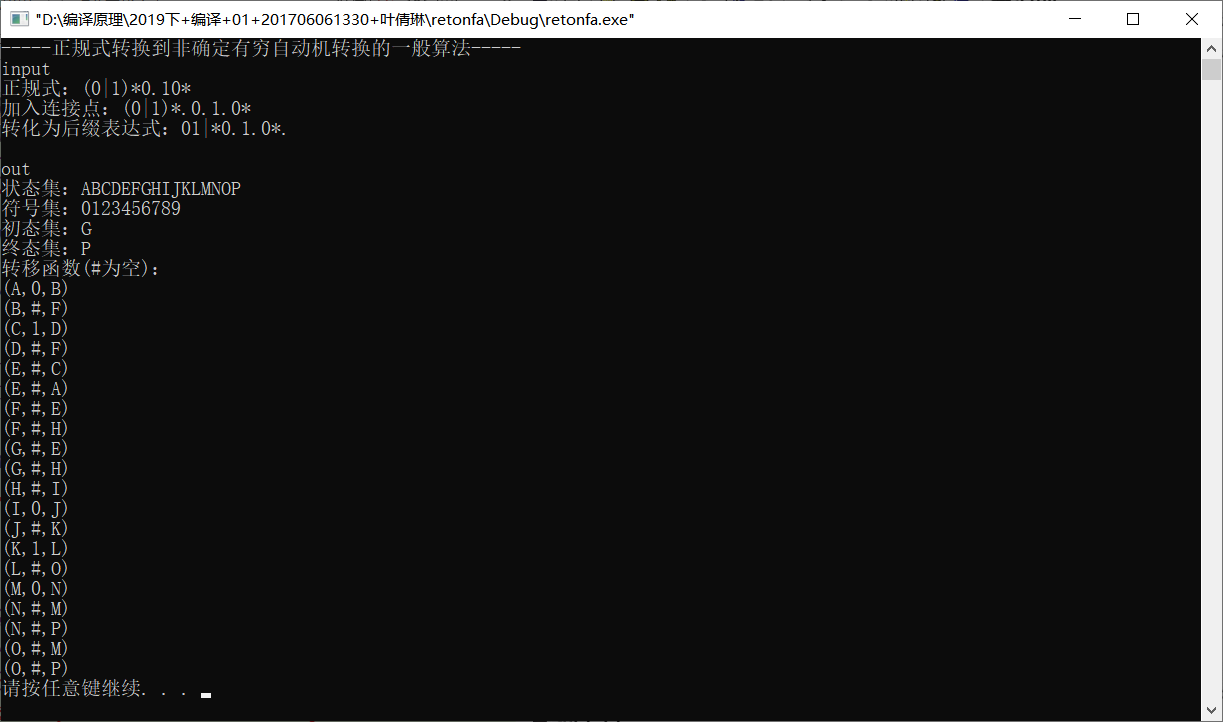


图4-1-1 正规式转NFA实验结果